

DOCUMENT RESUME

ED 259 475

EA 017 953

TITLE Teacher Compensation and Evaluation in Public Education.

INSTITUTION Indiana Univ., Bloomington. Center for Public Sector Labor Relations.

PUB DATE 85

NOTE 40p.; Funded in part by a grant from the Federal Mediation and Conciliation Service.

PUB TYPE Collected Works - Serials (022) -- Viewpoints (120)

JOURNAL CIT Dialogue: A Review of Labor-Management Cooperation in Public Education; v2 n1 Spr 1985

EDRS PRICE MF01/PC02 Plus Postage.

DESCRIPTORS Differentiated Staffs; Elementary Secondary Education; Evaluation Methods; *Incentives; Master Teachers; *Merit Pay; Peer Evaluation; Teacher Effectiveness; *Teacher Evaluation; Teacher Improvement; *Teacher Salaries; Teacher Supervision; *Teaching (Occupation)

ABSTRACT

This issue of "Dialogue," a review of labor-management cooperation in public education, focuses on the theme of teacher compensation and evaluation. After a brief introduction locating these topics within the context of nationwide educational reform, the first section describes the advantages and disadvantages of some of the better-known compensation and incentive approaches. These include the traditional single-salary schedule, along with newer, currently fashionable approaches such as performance-based pay, bonuses, and career ladders or differentiated staffing. The second section examines some of the factors related to the choice and use of an evaluation system. First, basic conceptions of teaching--as labor, craft, profession, or art--are reviewed as a basis for the subsequent discussion of the purposes of teacher evaluation. Thereafter, the timing and tools of evaluation are discussed, including teacher interviews, competency tests, indirect measures, classroom observation, student ratings, peer review, student achievement, faculty self-evaluation, and work samples. Finally, findings from a recent Rand study on the necessary conditions of effective teacher evaluation are discussed. References are included. (TE)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

CENTER FOR PUBLIC SECTOR LABOR RELATIONS

In response to the changing relationship between public employees and their governmental employers, the Center for Public Sector Labor Relations was established to provide an ongoing means through which practitioners, elected officials, and the public at large could exchange information and cooperatively solve problems related to public sector labor relations.

The Labor Center maintains a library on public sector labor relations and responds to specific requests for information. It has published the Midwest Monitor, a topical digest of current literature and developments, and the Practitioner Guide series of occasional monographs to help improve practitioners' labor relations skills. It currently publishes Dialogue, a quarterly review of labor-management cooperation in public education, which is available by individual subscription at \$8 per year.

Center for Public Sector Labor Relations
School of Public and Environmental Affairs
Indiana University
Bloomington, Indiana 47405
(812) 335 - 7851

Charles F. Bonser
Dean
School of Public and Environmental Affairs

Richard S. Rublin
Professor and Labor Center Director

TEACHER COMPENSATION AND EVALUATION
IN PUBLIC EDUCATION

dialogue is funded in part by a grant from the Federal Mediation and Conciliation Service. dialogue may not be reproduced, in whole or in part, in any form without express written permission from the Director. Annual subscriptions to dialogue are available for \$8. Subscription information is on the inside back cover. Address inquiries to the Director of the Labor Center.

INTRODUCTION

The reform of public education in the United States has begun to move beyond the rhetoric of the reports which provided the impetus for change. In the last few years, most states have enacted legislation and policies aimed at raising the quality of their education process and product. Change has been mandated in graduation requirements for high school students, teacher training, teacher compensation and a range of other elements in the education system. For example, as of January 1985:

- 43 states have raised high school graduation requirements; 5 are now considering more stringent ones
- 10 states have "honors" or "merit" diplomas requiring added and more rigorous courses
- 15 states require an exit test for graduation from high school; 4 more are considering similar requirements
- 37 states have instituted statewide student assessments at elementary and/or secondary levels
- 29 states have already upgraded teacher-education requirements to include mandated teacher competency tests; 10 more are considering the idea
- 28 states have made other changes in teacher certification procedures; 16 more are considering changes
- half the states now provide aid for prospective teachers (scholarships and/or forgivable loans)

- 12 states have approved across-the-board or minimum-level teacher salary increases; 17 more have proposals currently under legislative consideration
- 13 states have lengthened the school day or year; 7 more are considering doing so
- 13 states now limit student-teacher ratios
- 7 states have established pre-school programs; 8 more are considering them

These actions have been bolstered by increased education budgets in all 50 states during the past two years.¹ Other organizations and agencies have attempted to guide such state efforts by offering various policy recommendations. The Council of Chief State School Officers recently decided to support the development of a system of indicators of educational progress, even though it recognizes the potential of trivializing or narrowing academic outcomes. The Council will review a list of proposed indicators next November and presumably will disseminate its recommendations thereafter.² A similar approach was taken by the Department of Education's National Center for Educational Statistics, whose national task force has developed its own list of possible indicators of educational quality.³ The National Consortium for Educational Excellence, 32 school superintendents responding to the National Commission on Excellence in Education, has presented an agenda from those "on the firing line," incorporating in their recommendations the findings of local "commissions" and surveys. They offer both regulatory and developmental policy recommendations for public education pertaining to student learning, the quality of teachers and teaching, collaboration between schools, families and communities, and acceptance of responsibility.⁴

Joining the fray from a different perspective, the Public Education Information Network - a non-profit national organization of about 200 teachers, administrators, researchers, and teacher educators - urges "bottom-up" rather than "top-down" reform: "people in each school community must develop their own agendas and specific programs for reform."⁵ In support of maintaining local control of educational reform, the American Educational Studies Association recently reminded us of the importance of maintaining for local educators their freedom to innovate and make decisions. States and school districts, therefore, may wish to frame concepts and general curricula, but the content of instruction should be determined by teachers.⁶

The National Education Association, which represents the majority of organized teachers in United States, has also called for restructuring local responsibility for education in order to create greater flexibility for schools to serve students developing at different rates.⁷ The rival American Federation of Teachers, for its part, has called for the development of a nationally recognized licensing examination for new public and private school teachers.⁸ And teachers themselves "are wide open for change in reforming public schools, including those changes which will undoubtedly require sacrifices and extra effort from the teachers themselves," according to a recent national survey of United States teachers.⁹

Because teachers have received the majority of attention and much criticism in the reform movement, there is a need to understand clearly the issues that are involved in their competence and performance. This dialogue focuses on two of those inter-related factors: compensation and evaluation. While many other factors related to teachers in

particular and schooling in general must be considered in any holistic reform attempts, compensation and evaluation are clearly two of the most critical - and inextricable - and have received the most attention in the current spate of reports and reportage.

Teacher Compensation

The chain of assumptions undergirding the importance of teacher compensation may be stated roughly as follows: "Money (and the right management strategies) will influence who chooses teaching and who is motivated to stay in teaching; if highly qualified people choose to teach or remain in teaching, student outcomes will be affected positively; all school districts must fashion financial and other incentives that meet their own needs in order to attract and retain such qualified people." Unfortunately, there is not conclusive evidence that any compensation method will work to attract, retain, or motivate highly qualified teachers.¹⁰ Even so, there has been and will continue to be considerable debate over the best compensation method: single salary schedule, performance-based salaries, bonuses, market-sensitive salaries, salary differentiation, base salary, and benefit modifications - not to mention other financial incentives such as loan forgiveness, scholarships, grants, sabbaticals and training. Further complicating the compensation question are such related alternatives as career ladders and mentor programs that combine enhanced professional responsibility with compensation increments.

Following are descriptions of some of the more well-known compensation and incentive approaches, as well as some of the advantages and disadvantages of each.

Single Salary Schedule

By far the most prevalent form of compensation, the single salary schedule offers base pay and increments based on a teacher's education and experience (years teaching), regardless of performance in the classroom. While critics of this system relate it to the low scores of entering teachers on exams such as the National Teacher Examination and the Scholastic Aptitude Test (and thus, to the declining quality of schools), supporters point out that the single salary schedule was adopted widely before the decline in test scores began (the late 'sixties). Also, research linking teachers' standardized test scores and classroom effectiveness has been almost nil. Critics also argue that since pay advancement is not based on improvement in teachers' performance, there is no motivation for teachers to do a better job in the classroom. However, on this point, too, there is no direct evidence. Further, since the single salary schedule is based on the theory that teachers become more competent as they accumulate experience and education, they are, one could argue, receiving performance-based pay.¹¹

Regardless of the evidence of the relationship between the single salary schedule and teacher quality, motivation or retention, there seems to be considerable agreement among the major reports on education that teachers' base pay needs upgrading. The Carnegie Foundation for the Advancement of Teaching, The National Task Force on Education, The National Commission on Excellence in Education, The

National Consortium for Educational Excellence, among others, all call for raising teachers' salaries across the board, and more than one third of the states have done just that.¹² The fact of the matter is that teachers' salaries have lost ground to other professions and in real dollars. The annual rate of increase in teacher salaries nationwide, as of 1983, was nine percent.¹³ Teachers, of course, know their pay is low: nine out of ten say it is too low in their community, and almost the same number believes that low pay is the main reason teachers are leaving their jobs.¹⁴

Performance-based Pay

The most well-known form of performance-based pay is merit pay, which is clearly the approach garnering most of the education community headlines. It is based on the proposition that teachers vary in ability and/or motivation and thus their performance will differ. Linking performance and pay will (it is argued by proponents) motivate teachers to improve their performance. This approach assumes that teachers and administrators know what merit-worthy performance is, and that achieving merit-worthy performance is a matter of motivation and not some other factor(s), such as job training or job design.¹⁵

A recent paper by the National Education Association's director of research suggests that a true performance-based pay plan meets these criteria:

- roles, responsibilities, and tasks for management, teachers, and support staff are ranked according to skill and responsibility levels

- performance is ranked according to clear criteria related to roles, responsibilities, and tasks
- salaries are commensurate with the skills employed and responsibilities assumed and tasks performed¹⁶

In practice, merit pay plans in local education agencies exhibit a wide range of features and philosophies, and may be based less on cohesive assumptions than on political pressures. Plans labeled "merit pay" may offer one-time bonuses, or yearly increments on a basic salary schedule; they may use quotas, or give some recognition to all teachers; they may be based on observation of classroom performance, student test scores, accumulated training credits, assumption of extra duties, or a combination of some or all of these; they may focus on individual performance or team or school performance; they may have been in operation for decades or folded after a matter of months. An overview of merit pay approaches would show that

- 96% of local education agencies are not using pay for performance systems
- more LEAs that have tried merit pay systems have dropped them than retained them
- most LEAs invest little money in their merit pay systems
- the majority of merit pay plans use an assessment of input criteria to evaluate teachers (e.g., knowledge of subject, evidence of preparation)
- less than one in five tries to link merit pay to students' achievement test scores¹⁷

The actual structure of the pay scales is usually one of the following:

- multiple salary scales - determines the salaries of teachers in different performance groups (e.g., "average," "excellent," "superior")
- varying salary increments - for differing performance assessments
- accelerated increment rate - permits more rapid movement toward the top of the salary scale (e.g., double increment per year)
- exceeding the maximum salary step - allows outstanding teachers to move beyond the top of the salary scale¹⁸

It is worth noting that public school teachers themselves have mixed feelings about merit pay. A recent survey found that 71% of the teachers responding believe that merit pay could work if a teacher's merit can be judged on an objective standard. However, 59% of the respondents disagreed that merit pay is an effective way of attracting and retaining good teachers in the profession.¹⁹

Where local districts have tried and dropped merit pay systems (and there is not a wealth of information on such districts), administrators have noted these problems:

- unsatisfactory evaluation procedures - the key difficulty
- administrative problems (e.g., changes in school leadership, time needed for evaluation, and record-keeping)
- staff dissension - lower teacher morale, union opposition, etc.
- financial problems - lack of funds to support the system or permit clear distinction in meritorious vs. non-meritorious performance, cost to local taxpayers²⁰

An excellent summary of the possible advantages and disadvantages of performance-based salaries has been prepared for a joint report of three professional education association, and it is reproduced below.²¹

Possible Advantages	Possible Disadvantages
- Opportunity to earn higher salary for superior performance may attract higher quality candidates. - Monetary rewards for performance may improve morale and retention of superior teachers. - Rewarded teachers are likely to be more highly motivated. - Additional expenditures may be seen as justified if results are substantial. - If fully implemented, adequately financed, and accepted by teachers, it may have among the strongest effects of all incentive plans. - An effective performance appraisal system is required, which may support improvement and provide a basis for dismissing incompetent teachers. - Teachers are rewarded for concentration on teaching; additional roles and responsibilities are not necessary. - Some versions may be easier to administer than more complex structural changes. - If appraisal system is sound, decision making is straightforward.	- Some strong candidates and existing staff may dislike competitiveness of plan. - Morale of unrewarded teachers may decline. - Cost may be high if quotas are avoided, many teachers qualify, and rewards are sufficiently large to be seen as worthwhile. - Performance appraisal may be expensive. - Negative effects may outweigh positive if rewards are insufficient or if destructive competition and jealousy result. - If insufficient time and resources are devoted to developing a strong appraisal system, the incentive plan may lose credibility and support. - Development of an effective appraisal system is difficult. - Significant attitudinal changes are required for acceptance. - Emphasis on performance represents a large shift in many districts. - Management of complex appraisal system can be demanding; such time is required.

Possible Advantages	Possible Disadvantages
- Community will be supportive if plan attracts and retains better teachers and improves outcomes. - Teachers will be supportive if plan accurately discriminates among levels of performance.	- Community may be opposed if it appears that more is being spent for the same people and results. - Teachers will be opposed if decisions are subjective, criteria are too narrow, and competition replaces collegiality.

Bonuses

Some "merit pay" systems are actually bonus plans. Rather than adding salary increments, a bonus is a one-time award. Bonuses serve a variety of purposes, not all of which directly have to do with classroom performance. They may be used to induce college graduates to enter teaching, to reward annual performance, or to attract teachers to teach in "difficult" or less attractive schools. An example of the extensive use of bonuses is Houston's Second Mile Plan, wherein teachers may earn bonuses for

- teaching in schools with concentrations of disadvantaged students
- teaching in areas of critical staff shortage
- exceeding the qualifying attendance requirements
- successfully completing courses in subjects relevant to a teacher's current assignment
- serving in schools that exhibit outstanding educational progress
- serving in "unique campus assignments"

To qualify for bonuses, teachers must hold a valid certificate, receive an acceptable assessment, demonstrate proficiency in academic skills, and have

five or fewer annual days of absence (among other criteria).²²

In the Dallas Independent School District, bonuses may be awarded to all staff members in the 25% of the schools with the best achievement test gains relative to projected gains. In Seiling, Oklahoma, teachers may receive different bonuses for achieving individual and school objectives.²³

The following table shows some possible advantages and disadvantages of bonuses.²⁴

Possible Advantages	Possible Disadvantages
- Bonuses may be effective in attracting strong candidates in areas of shortage or especially qualified persons in all areas. - Bonuses may encourage superior teachers to remain if used to reward performance. - Motivation to perform well is renewed annually. - Negative effects on nonrecipients' morale may be avoided by keeping information confidential. - Teachers can be encouraged to attempt needed improvement. - Bonuses do not affect base salaries in succeeding years. - Bonuses do not affect retirement costs. - Bonuses can have substantial effects if they are substantial. - Bonuses for goals other than performance (such as entering teaching, teaching in certain schools, or improving attendance) do not require change in performance appraisal systems.	- Bonuses may be less effective than higher salaries in attracting strong candidates, who may be concerned that bonuses will not last beyond the initial year(s). - Some of the best performances may go unrewarded if program is voluntary and some of the best teachers decline to participate. - Funds may be more vulnerable to budget cutting than salary funds. - Cost may be high if bonuses are sufficiently numerous and large to motivate desired behavior. - Positive effects may be limited to particular teachers or objectives if program is too narrow. - If bonuses are not related to performance, the most important accomplishments may go unrewarded.

Possible Advantages	Possible Disadvantages
- If bonuses are tied to performance, impetus may be given to improvements in the appraisal system. - Bonuses may be implemented without organizational or other fundamental changes. - For non-performance goals, criteria will be objective and decision making will be straightforward. - Increased income may lead to teacher support. - Voluntary participation and confidentiality may make plan more acceptable to teachers. - Community will be supportive if results are improved and if costs can be restricted by not affecting salaries in subsequent years and retirement demands.	- If bonuses are tied to performance, the difficulties inherent in developing an effective appraisal system must be overcome or choices will lack credibility. - Bonuses are not a familiar concept in education and may take time to be accepted. - If performance is to be rewarded, substantial time will be required. - If bonuses are used for various purposes, plan may become complex. - Teachers may oppose if they believe that the wrong goals are rewarded, some teachers have an advantage over others, or decisions are subjective. - Community will oppose if it believes that more money is being spent without visible improvement.

Career Ladders/Differentiated Staffing

Although research and reports attempt to distinguish between career ladders and differentiated staffing, the practical differences are unclear. A career ladder is a modification of the traditional structure of the teaching career, while differentiated staffing increases teachers' compensation and makes the job more interesting by extending and varying teachers' responsibilities.²⁵ In any case, the rungs of the ladder are ranks that require the use of more advanced skills as the teacher assumes increasingly complex and demanding duties. Bottom rungs typically include more clerical and supervi-

sory duties not central to teaching, while the top rungs often include tasks such as curriculum development and program coordination.²⁶

These approaches are viewed as ways of retaining quality teachers, reducing turnover and placing the effects of turnover at the lower end of the career ladder, thus generating a more highly-qualified and specialized staff, dedicated to professional growth, and thereby enhancing the overall quality of education in public schools.²⁷

There is some question, however, of the effectiveness of such plans. They may only institutionalize what they are aimed at correcting, since the most qualified and ambitious teachers spend far less time in teaching than in administration. In effect, good teachers are forced to assume nonteaching tasks if they want to reach optimum salary levels. Further, if the system really works (i.e., if it keeps good teachers, who reach the top of their pay ladders), teachers at lower levels have little hope for promotion. The lack of mobility would be a disincentive for those wanting greater responsibility and pay, and they might become discouraged and leave the school system or the profession. And perhaps the most difficult question to answer about these plans is, where do newly created positions to "build" the ladder or staffing pattern come from? Current classroom duties might well become diluted or "simplified," and thus the initial excitement and challenge of a new position (lower rungs are usually filled by new/younger teachers) are likely to be dissipated.²⁸

The typical career ladder has three or four levels:

- I. entry/probation/apprentice
- II. continuing/regular/career

III. senior

IV. mentor/master

All new teachers enter at the first level, as usually must experienced teachers from other districts or systems. Probationary periods will vary in length, leading to the next step after some combination of performance, years of service, and educational attainment criteria are met. New titles usually accompany advancement. The upper levels usually allow/require observation and/or training of other teachers, leading instructional programs, or evaluating other teachers. Contract extensions to 11 or 12 months may be offered.²⁹

Past efforts at establishing career ladders or differentiated staffing plans have encountered a range of problems:

- hasty implementation resulting in miniscule differences among the various roles in the plans
- complicated funding arrangements resulting in cost overruns
- lack of appeal to teachers, who were mostly women who did not wish to relinquish the ground toward equity they had won in the single salary schedule
- resistance of the organized teaching profession
- lack of continuing federal and state support that would see the innovation through its rough early stages
- top-down patterns of implementation resulting in weak teacher loyalty
- abandonment of plans when leadership changed
- inability of the models to demonstrate they had contributed to improved learning³⁰

Career ladders and differentiated staffing are an important part of the efforts of state legislatures to upgrade education in their school systems. Currently, at least 25 states have adopted, or are trying, deliberating, developing or proposing such plans.³¹ Teachers, too, seem to favor the option, as indicated by a recent survey wherein 87% of teacher respondents felt positively about establishing career ladders that provide teachers with greater opportunities to take on more responsibility and pay.³² Although only a few of the states have actually implemented their plans (e.g., Tennessee, Florida) and a few local school systems (e.g., Charlotte-Mecklenburg), there is a clear push forward nationally, regardless of previous research being inconclusive or even discouraging. It is worthwhile considering the advantages and disadvantages of this approach in summary form, as presented in the following table.³³

Possible Advantages	Possible Disadvantages
- Opportunities to advance to higher levels within teaching may encourage superior teachers to remain, and may attract strong candidates who believe that they have potential for advancement. - Superior teachers should be motivated by opportunities for advancement and variation in activities. - Initial investment is reasonable. - Additional resources reflect merit and support additional activities. - Because part of reward is prestige, secretary requirements are lessened.	- Teachers who do not advance to higher levels may become discouraged. - Potential teachers may be less attracted by long-term possibilities than by short-term rewards. - Workload may increase, lowering morale. - Efforts may be directed away from the classroom. - Morale may decline among teachers who do not advance. - Long-term commitment is required. - If many teachers advance, if additional responsibilities are many, or if rewards are large, the system may be costly. - Available funds may be directed away from base salaries.

Possible Advantages	Possible Disadvantages
- Motivation and retention may be substantially improved if fully implemented. - Structure lends itself to long-term, multi-faceted appraisal, which may have greater reliability, validity, and acceptance than annual appraisals. - It may be one of the most effective ways of achieving the substantial changes required of fundamental improvement. - Add-on feature requires little fundamental change in district operations. - Community can see more work for more money. - Teachers may support emphasis on teacher development and longer-term performance appraisal. - Teachers may appreciate change in activities.	- If additional responsibilities are overemphasized, effects may be outside of core teaching and learning goals. - It is difficult to obtain short-term results. - Initial implementation difficulties may exist. - It requires considerable organizational change. - Expanded support programs may require substantial administrative work in development and execution. - Teachers and community may not support use of time and funds for nonclassroom activities. - Shorter-term results may be desired.

Teacher Evaluation

The selection and successful implementation of a performance-based or career ladder-type teacher compensation system must be accompanied by some form of teacher evaluation. Much of the controversy surrounding the adoption of new teacher compensation systems centers on the validity, reliability, objectivity, and acceptability of the evaluation system from which data are derived to justify decisions about who gets paid for various levels of merit. This section of dialogue examines some of the factors related to the choice and use of an evaluation system.

Conceptions of Teaching

Different conceptions of teaching suggest different methods of evaluation. There are four ways of looking at teaching that tend to predominate: teaching as labor, craft, profession, or art. Darling-Hammond et al.³⁴ describe each and its associated evaluation strategy. Their discussion is summarized here.

Teaching as labor - Teaching is rationally planned, programmatically organized, and routinized in the form of standard operating procedures by administrators; teachers implement and adhere to specifications; evaluation involves direct inspection of the teacher's work (lesson plans, classroom performance, performance results).

Teaching as craft - Teaching requires a repertoire of specialized techniques and generalized rules for their application; teachers carry out assignments without detailed instructions or close supervision; evaluation is indirect and involves determining that the teacher has the requisite skills.

Teaching as profession - Teaching not only requires a repertoire of specialized techniques but also exercise of judgment about when to apply those techniques, so the teacher must also master a body of theoretical knowledge; evaluation standards are developed by peers, and evaluation focuses on teacher competence at professional problem-solving.

Teaching as art - Teaching techniques and their application may be novel, unconventional or unpredictable, and standards of practice are personalized - teaching uses science but is not itself a science; evaluation involves both self-assessment and critical assessment by others, entailing the study of

holistic qualities rather than analytically derived quantities - it relies on high-inference variables and observation of patterns of events (rather than low-inference variables and counts of specific, discrete behaviors).

Although these conceptions are rarely found in pure form, their differences signal the need for attention to the philosophy of the system or individuals establishing or conducting teacher evaluation. To the extent that clarity can be obtained on conceptions of teaching, appropriate evaluation strategies can be identified.

Purposes of Teacher Evaluation

It is also important to be clear on the purposes for which teacher evaluation is conducted. A draft of the recent Rand study³⁵ provides a concise statement of basic purposes of teacher evaluation:

<u>Purpose/Level</u>	<u>Individual</u>	<u>Organizational</u>
Improvement	individual staff development	school improvement
Accountability	individual personnel (job status decisions)	school status (e.g., certification)

- In general, teacher evaluation processes most suited to accountability purposes must be capable of yielding fairly objective, standardized and externally defensible information about teacher performance.

- Evaluation processes useful for improvement objectives must yield rich, descriptive information that illuminates sources of difficulty as well as viable courses for change.
- Evaluation methods to inform organization decisions must be hierarchically administered and controlled to ensure credibility and uniformity.
- Evaluation methods to assist decisions about individuals must consider the context in which individual performance occurs to ensure appropriateness and sufficiency of data.

Another way of looking at evaluation purposes is through the formative/summative dichotomy. Formative evaluation is concerned with helping to improve a development or implementation process, while summative evaluation is concerned with evaluating a program or process after it is in operation or after it has been completed. Formative evaluation is a collaborative process, involving the evaluator in the improvement of the process (or person) being evaluated, while summative evaluation is independently conducted and should be technically excellent research assessing the impact of the process or program under inspection. Given these differences, evaluation systems that allow teachers to improve their performance are formative, and systems that reward and/or punish teachers are summative. Research and experience have shown that it is unwise to assign both formative and summative responsibilities to one evaluator, and it is almost impossible to have a teacher evaluation system that successfully provides reward and/or punishment and improves teaching performance. This suggests a strategy wherein the initial purpose of a teacher evaluation system is formative - helping teachers to improve their performance, if needed - and if there are some teachers for whom the formative/improvement

process does not work, these teachers can be switched over to a second, distinct, summative system. A summary³⁶ of formative/summative philosophy, theory and practice follows:

Formative Evaluation

Summative Evaluation

Philosophy:

Each individual strives for excellence.

Individuals achieve excellence only if supervised or evaluated by others.

Theory:

Evaluations done to improve the performance of the individual. Reward or punishment should be decided internally.

Evaluation is done to improve the performance of the social system. Reward or punishment should be decided externally.

Practice:

Evaluate the process of instruction but not the person.

Evaluate the products of instruction as well as the process and the person.

Timing and Tools for Teacher Evaluation

Teacher evaluation may be conducted at any of five points in a teacher's career: (1) on entry into preservice preparation; (2) on completion of preservice education and before entry into practice; (3) on the basis of the teacher's behavior while

teaching; (4) on the basis of the pupils' behavior during teaching; (5) on the basis of learning outcomes. At the first two points, evaluation is of personal qualities which are retained in any situation in which the teacher may be employed. Evaluation at the last two points is based on pupil (rather than teacher) behavior, and does not measure anything about the teacher directly. It is at the third point that teacher qualities may affect pupil behavior, the point of teacher performance.³⁷

Although teacher performance has been the traditional focus of evaluation, other evaluation points are receiving increased attention in recent years. The following outlines evaluation processes and methods that are applicable at various points in a teacher's career.³⁸

Teacher Interviews. Interviews (or conferences) have been common evaluation tools, and continue to be part of more recent approaches. They are used for hiring decisions and to communicate performance appraisals to teachers. Although they have been generally informal and unstructured, there are some standardized methods (e.g., Teacher Perceiver Interview). Preobservation conferences that involve teachers in their own professional development may have some negative aspects in locking teachers into a predetermined style that inhibits individual creativity in the classroom.

Competency Tests. There will soon be a great increase in the use of standardized tests for initial certification and hiring. State legislatures and local school districts are urging adoption of such tests, believing that teachers should be able to show cognitive competence as a teaching prerequisite. Tests like the National Teacher Examination (NTE) measure academic achievement and

preparation in areas of general education, professional education and specialized subject areas. However, there is no consistent relationship in the research literature between NTE scores and measures of teacher performance. Advocates of competency tests believe that the tests guarantee a minimum level of knowledge, eliminate interviewer bias, and are legally defensible.

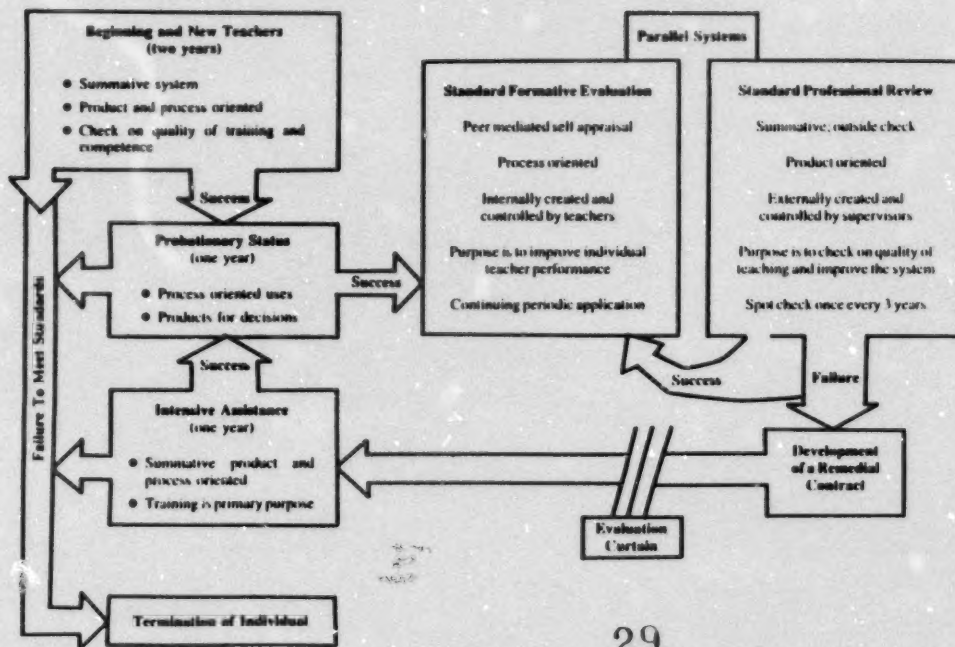
Indirect Measures. Measures of teacher skills, attitudes, interests, and abilities do not seem to discriminate between effective and ineffective teachers, yet they are recommended as supplementary sources of evaluation data by some researchers. In particular, teacher flexibility or adaptability may be linked to effectiveness. Some characteristics may be more effective in some teaching contexts than others.

Classroom Observation. Together with teacher interviews, classroom observation is the main tool for teacher evaluation. It can result in measures of performance that reveal a view of the climate, rapport, interaction and functioning of the classroom available from no other source. The actual techniques involved vary greatly in structure and method. Most often the principal acts as observer, although trained evaluators, school administrators, other teachers and students may observe or rate the teacher. Time and school board policies regarding evaluation may constrain the frequency and length of periods of observation. The advantage of seeing a teacher in action is balanced (or outweighed, some say) by such limitations as observer bias, insufficient sampling of performance, and poor measurement instruments. Interrater reliability and validity have been lacking in supervisory ratings, and ratings have had limited stability and generalizability, especially when low-inference measures are used.

Student Ratings. As another form of classroom observation, student ratings may be useful in secondary, and perhaps elementary, schools if (a) students know when they have been motivated; (b) it is the student whose behavior is to be changed; (c) student ratings are used as feedback to the teacher; and (d) student recognition may motivate good teaching. Although student ratings are inexpensive and highly reliable (.8 to .9 or higher), and some studies have found modest correlations with student achievement, their validity and utility are questioned.

Peer Review. This consists of examining lesson plans, exams, examples of graded exams, and classroom observation. It looks at both classroom performance and what the teacher intends to have happen, as well as other behavior as exhibited by assignment and grading practices. This technique assumes that peers are in the position to assess competence and also can offer specific and practical suggestions for improvement. Because it is more open to divergent criteria for assessing performance and is not subject to direct administrative control, it is not generally recommended for use as the basis of personnel decisions. However, when incorporated as part of a larger system of evaluation, peer review can address many of the concerns of teachers and administrators for an approach that clearly focuses on the formative/improvement aspect of teacher development. Such an approach is presented in the model displayed on the following page.³⁹

Student Achievement. Many people believe that the ultimate test of a teacher's effectiveness is student achievement, and that student performance on standardized achievement tests can be used - in a variety of ways - as a measure of how good a teacher is. They urge comparing student test scores to a



national norm, comparing test score gains with those of a comparable class, assessing net gain over time, comparing achievement scores to aptitude scores (usually of whole classes, and validating by inspecting scores of classes taught over more than one testing period), and so forth. Unfortunately, these "logical" procedures have some serious problems.

Studies of student test scores as a measure of teaching effectiveness consistently indicate that reliability is quite low; that is, the same teacher produces markedly different results in different situations.⁴⁰ There are many influences on students' test scores other than teacher performance (e.g., school and class size; quantity and quality of available instructional resources, materials and support personnel; centralization of school decision-making; school discipline practices; peer group influences; curriculum and textbook decisions; and the ways in which schools are organized, including the amount of time allocated for instruction and the ways in which students are grouped into classes). Another problem is that teachers work with classes of differing achievement levels, and these differences cannot be fully overcome by statistical adjustments. Also, teaching effects are both long- and short-range, and they may show up in different student learning behaviors at different points.⁴¹

But there are several even more basic and serious problems with the use of student scores on standardized achievement tests. Such tests measure a very few things fairly well, a large number of things pretty poorly, and some extremely important things not at all. They are arbitrary samples, to some extent, of large bodies of knowledge, and items are excluded which are answered correctly by too many, too few, or the wrong subset of students.

They do not tell what a child knows - only how a child ranks against peers. By reducing a complex activity to its tiny discrete components, they measure poorly how well students can apply knowledge and skills to real-life tasks. Further, they tend to foster standardized teaching, teaching to the test, and learning via a single type of thinking and a single mode of performance.⁴²

Although achievement tests can provide useful diagnostic information to tell teachers how their students are doing on certain tasks, many teachers are not adept at constructing and interpreting tests for such purposes.⁴³

Faculty Self-evaluation. When combined with individual goal-setting, faculty self-evaluation may promote self-reflection and motivation toward growth. Teachers can use a variety of data (e.g., student or peer ratings, self-assessment measures of student achievement, etc.) to judge their own teaching. Externally developed data may allow teachers to assess their strengths and weaknesses against both personal and organizational standards. This method is not suitable for accountability decisions, but can be used for staff development.⁴⁴

Work Samples. Teacher products represent what a teacher is trying to accomplish, and also can be evidence of professional development. A teacher file might include a rationale for each course or year's methods and coverage; evidence about such professional development activities as readings, courses, workshops and consultation; a description of teacher innovations and experiments; and a list of courses taught, enrollments, grades, extra curricular duties, committee assignments, and other formal service.⁴⁵

Necessary Conditions

The recent Rand study, while not intended to generate a recipe for successful teacher evaluation, does offer conclusions and recommendations that may constitute a set of necessary but perhaps not sufficient conditions. It suggests that the sufficient condition is leadership. The necessary conditions are set forth below.⁴⁶

A teacher evaluation system is more likely to succeed - to achieve its objectives - if it is compatible with a school district's educational goals, management style, conception of teaching work, and dominant community values.

- A school district should not adopt a teacher evaluation system simply because it has been found to work in other districts or simply because it is in vogue.
- A school district should examine its educational goals, management style, conception of teaching work and dominant community values before choosing a teacher evaluation system.

Commitment and resources for evaluation are far more important than checklists and procedures.

- Evaluators must be given sufficient time, unfettered by competing administrative demands, for evaluation activities. This may mean assigning staff other than the school principal to some evaluation functions.
- The quality of evaluation activities must be regularly inspected by higher authorities. The outcome of these assessments should be feedback to individual evaluators and input into the continuing evaluator training process.

- The oversight of teacher evaluation must include assessment of individual and collective evaluator competence and a means of continually addressing training needs.

It is important to decide what the main purpose of teacher evaluation will be and to match the process to the purpose.

- A school district should examine its motives for establishing teacher evaluation and prioritize them.
- A school district should examine its existing teacher evaluation system and decide which, if any, purpose it serves well. Changing the purpose should mean changing the process.
- A school system should decide whether it can afford more than one teacher evaluation process or whether it must choose the single process most applicable to its current needs.

For resource commitments and political support to be sustained, the benefits of evaluation must exceed the costs. The cost-effectiveness (or utility) of an evaluation process depends on the efficient deployment of a critical mass of resources.

- Resources must be commensurate with the numbers of teachers to be evaluated and with the importance and visibility of evaluation outcomes.
- Resources should be appropriately targeted so that real benefits occur; the process should not be diluted by inadequate time or lack of expertise.

Teacher involvement and responsibility improves the quality of her/his evaluation.

- Expert teachers should be involved in the supervision and assistance of their peers, particularly beginning teachers and those in need of special assistance.
- Teacher organizations should be involved in the design and continuing oversight of teacher evaluation to ensure its legitimacy, fairness and effectiveness.

At a minimum, a teacher evaluation system should involve a school system's staff in the following ways.

- All actors in the system have a shared understanding of the criteria and processes for teacher evaluation;
- All actors understand how these criteria and processes relate to the dominant symbols of the organization, that is, there is a shared sense that they capture the most important aspects of teaching, that the evaluation system is consonant with educational goals and conceptions of teaching work;
- Teachers perceive that the evaluation procedure enables and motivates them to improve their performance; and principals perceive that the procedure enables them to provide instructional leadership;
- All actors in the system perceive that the evaluation procedure allows them to strike a balance "between adaptation and adaptability, between stability to handle present demands and flexibility to handle unanticipated demand"; that is, that the procedure achieves a balance between control and autonomy for the various actors in the system.

* * * * *

The preceding discussion of teacher evaluation relies on considerable research and theoretical discussion. Unfortunately, it is not well-informed by studies of actual school practice. In preparation for such a study in southern Indiana school systems, researchers recently reviewed literally thousands of references to teacher evaluation and related areas. Among all these, less than ten studies examined actual school practice, none very extensively.⁴⁷ This lack of research underlines the caution which advocates of any particular teacher evaluation methodology would be well advised to adopt. It is all well and good to suggest that teacher evaluation can yield a range of benefits: improved administrator-teacher communication, improved consensus on instructional and school goals, enhanced mastery and confidence among teachers, a firmer basis for personnel actions, increased public confidence in schools, and better district-teacher union relations.⁴⁸ However, considerable care must be taken to avoid the rapid adoption-rapid disappointment syndrome that occurs all too often with innovations in the schools.

REFERENCES

1. Bridgman, Anne. "States Launching Barrage of Initiatives, Survey Finds." Education Week, February 6, 1985.
2. Council of Chief State School Officers. "Education Evaluation and Assessment in the United States." November 1984.
3. National Center for Educational Statistics. Indicators of Education Status and Trends. Department of Education, January 1985.

4. National Consortium for Educational Excellence. "An Agenda for Educational Renewal: A View from the Firing Line." Peabody College of Vanderbilt University, 1985.
5. Public Education Information Network. "Education for a Democratic Future. Equity and Excellence: Toward an Agenda for School Reform." 1985.
6. Rayvid, Mary Ann, Tesconi, Charles A., Jr., and Warren, Donald R. Pride and Promise. Schools of Excellence For All the People. American Educational Studies Association. 1984.
7. National Education Association. "An Open Letter to America on Schools, Students, and Tomorrow." (no date)
8. Jacobson, Robert L. "AFT's President Proposes National Exam for Teachers Akin to Tests for Professionals." Chronicle of Higher Education. February 6, 1985.
9. Louis Harris and Associates, Inc. The Metropolitan Life Survey of The American Teacher. June 1984. Metropolitan Life Insurance Company. New York.
10. Bacharach, Samuel B. et al. Teacher Compensation Systems and the Quality of Education. Organizational Analysis and Practice, Inc. Ithaca, New York. (no date), p. 9.
11. Bacharach, p. 16-19.
12. Bridgman, February 6, 1985.
13. Cresap, McCormick and Paget. Teacher Incentives. A Tool for Effective Management. Na-

tional Association of Secondary School Principals, Reston, Va., 1984, p. 8.

14. Gallup, Alec. "The Gallup Poll of Teachers' Attitudes toward the Public Schools." Phi Delta Kappan, October 1984, p. 98.
15. Bacharach, p. 30.
16. Masters, W. Frank, Jr. "Changing pay systems for teachers: key ideas and questions to be answered." For the NEA Blue Ribbon Task Force on Educational Excellence. October 1983.
17. Astuto, Terry A. and Clark, David L. Merit Pay for Teachers. Educational Policy Studies Series, No. 1. School of Education, Indiana University, January, 1985, p. 2.
18. Johnson, Susan Moore. "Merit Pay for Teachers." Harvard University, December 14, 1983.
19. Metropolitan Life Survey, p. 43.
20. Astuto and Clark, p. 6.
21. Cresap, McCormick and Paget, p. 18.
22. Cresap, McCormick and Paget, p. 19.
23. Cresap, McCormick and Paget, p. 19.
24. Cresap, McCormick and Paget, p. 20.
25. Astuto and Clark, p. 30.
26. Masters, p. 2.
27. Bacharach, p. 70.

28. Bacharach, pp. 72-76.
29. Cresap, McCormick and Paget, pp. 22-23.
30. ASCD Task Force on Merit Pay and Career Ladders. "Incentives for Excellence in America's Schools." The Association for Supervision and Curriculum Development, 1985, pp. 8-9.
31. Astuto and Clark, pp. 28-29.
32. Louis Harris and Associates, Inc., p. 40.
33. Cresap, McCormick and Paget, p. 24.
34. Darling-Hammond, Linda, Wise, Arthur E., and Pease, Sara R. "Teacher Evaluation in the Organizational Context: A Review of the Literature." Review of Educational Research, Fall 1983, v. 53, No. 3, pp. 285-328.
35. Wise, Arthur E., Darling-Hammond, Linda, McLaughlin, Milbrey W. and Bernstein, Harriet T. "Teacher Evaluation: A National Study of Effective Practices." National Institute of Education, December 1983 (Working Draft), pp. 13-15.
36. Barber, Larry. "Teacher Evaluation and Merit Pay." Education Commission of the States, Denver, Colorado. (no date) pp. 7-8.
37. Medley, Donald, Coker, Hower, and Soar, Robert. Measurement-Based Evaluation of Teacher Performance. Longman: New York, 1984.
38. Darling-Hammond et al, pp. 304-308.
39. Barber, p. 36.

40. Darling-Hammond et al., p. 307.
41. Darling-Hammond, Linda. "Taking the Measure of Excellence." American Educator. Fall 1984, pp. 26, 28.
42. Darling-Hammond, pp. 28-29.
43. Cresap, McCormick and Paget, p. 33.
44. Darling-Hammond et al., p. 308.
45. Cresap, McCormick and Paget, p. 34.
46. Wise et al. Ch. VI.
47. Smith, Gerald and Walden, James. "A Study of Teacher Evaluation Practices." Indiana University, School of Education, November 1984, pp. 2-3.
48. Cresap, McCormick and Paget, p. 30.

DIALOGUE

Subscription Form

Please enter my subscription to dialogue at \$8 for a full year (4 issues) .

_____ Number of Subscription (at \$8/year)

_____ Payment Enclosed

_____ Please Bill Me

_____ Please Bill My Agency (enclose purchase order where required)

(Make checks payable to : Center for Public Sector Labor Relations)

Please check mailing address on reverse side and make necessary changes :

Name : _____

Title : _____

Organization : _____

Address : _____

City : _____

State : _____ Zip _____

Detach this page and mail to :

Center for Public Sector Labor Relations
Indiana University
School of Public and Environmental Affairs
Bloomington, IN. 47405